

ABSTRACT

A method and a carrier medium carrying code segments to cause a processor to implement a method for resolving a possibly incorrectly entered URL. The method includes accepting the entered URL, parsing the accepted URL into URL parts, and carrying out a conventional URL lookup. In one embodiment, for any part of the accepted URL that is not valid, the method includes determining a signature for the accepted URL part; and conducting a fuzzy search for at least one valid URL part that is close to the invalid URL part according to a distance measure that combines at least one local measure, each measure suited for a particular type of URL part. At least one valid URL may be formed from the URL parts found in the fuzzy search. In one implementation, the conducting of the fuzzy search includes: determining at least one cluster of a set of pre-formed clusters wherein the accepted URL part is likely to be. Each cluster includes a set of valid URL parts that are close according to a distance measure, and has a representative URL part having a known signature. The determining of the cluster(s) includes finding the at least one signature of representative URLs close to the signature of the accepted URL part. The method includes further searching for a valid URL part within the at least one determined cluster.